

> O <
OI IO IntelliGenetics
> O <
FastDB - Fast Pairwise Comparison of Sequences
Release 5.4
Results file us-08-877-150b-35-inv.res made by sdevaid on Wed 24 Oct 101 13:42:24-PDT.
Query sequence being compared:US-08-877-150B-35' (1-2810)
Number of sequences searched: 1
Number of scores above cutoff: 1
Results of the initial comparison of US-08-877-150B-35' (1-2810) with:
File : u49853.seq

100-
N
U
50-
-
-
-
R
-
O
F 10-
S
E 5-
Q
U
E
N
C
E
S
SCORE 0 11 22 33 44 54 65 76 87 98
STDEV

PARAMETERS

Similarity matrix Unitary K-tuple 4
Mismatch penalty 1 Joining penalty 30
Gap penalty 1.00 Window size 32
Gap size penalty 0.33
Cutoff score 0
Domination group 0

SEARCH STATISTICS

Scores: Mean Median Standard Deviation
98 0 0.00
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
Number of residues: 1529
Number of sequences searched: 1
Number of scores above cutoff: 1

The scores below are sorted by initial score.
Significance is calculated based on initial score.
A 100% identical sequence to the query sequence was not found.
The list of best scores is:

Sequence Name	Description	Length	Score	Init. Opt.
1. u49853	TOIG of: u49853 check: 2626 from: 1 to: 1529	1529	98	660 0.00 0
1. US-08-877-150B-35' (1-2810) u49853 TOIG of: u49853 check: 2626 from: 1 to: 1529 TOIG of: u49853 check: 2626 from: 1 to: 1529 LOCUS MMPTPHSC 1529 bp mRNA ROD 08-NOV-1996 DEFINITION Mus musculus protein tyrosine phosphatase mRNA, complete cds. ACCESSION U49853 VERSION U49853.1 GI:1293621 KEYWORDS house mouse. SOURCE Mus musculus ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. REFERENCE 1 (bases 1 to 1529) AUTHORS Cheng, J., Daimaru, L., Fennie, C. and Lasky, L.A. TITLE A novel protein tyrosine phosphatase expressed in JOURNAL Blood 88 (4), 1156-1167 (1996) MEDLINE 96329547 REFERENCE 2 (bases 1 to 1529) AUTHORS Cheng, J., Fennie, C. and Lasky, L.A. TITLE Direct Submission JOURNAL Submitted (23-FEB-1996) Laurence A. Lasky, Molecular Oncology, Genentech Inc., 460 Pt. San Bruno Blvd., San Francisco, CA 94080, USA COMMENT On May 2, 1996 this sequence version replaced gi:1277080. FEATURES Location/Qualifiers 1..1529 /organism="Mus musculus" /db_xref="taxon:10090" /cell_line="hematopoietic progenitor cell" 19..1380 /codon_start=1 /product="tyrosine phosphatase" /protein_id="AAB18623.1" /db_xref="GI:1293622" /translation="MSRHTDLVRSFLQLEARDYREGAILAREFSIDIKARSVAMKSEG VCSTKAGSLGNTNRYKDVAYDETRVILSLQEEHGDIYINAFRLGDSQAYI ATOGPLHLLDFWRLWVEGVKVLMAQETENGRRKRCERYWAEQPLAGPCIT LTRETLNADITLTLQVTFQKESVHQVMSPDHGVSSSDHILTMVEEARCLQ GLGPGCLVHCSCAGRTVLCAYDYVRQLLTQIPNFSLFQVLMRKQRAAVQ TEQYRFLYHTVAQLFSRTLDQTSPIHYQNLKENCAPICKEAFSLRTSALPAPPG GVLSISVPAPPTLPMDATYAVQKRGASAGTGPCRAPSTDTPIYSQVAPRAQRPV AHTEDAQGTALRRVPADQNSGPDAYEEVTDGATGGLGNLRIGRPGKPRDPPAEW TRV" BASE COUNT 356 a 442 c 446 g 285 t ORIGIN U49853 Length: 1529 October 24, 2001 13:28 Type: N Check: 2626 .. Initial Score = 98 Optimized Score = 660 Significance = 0.00 Residue Identity = 49% Matches = 798 Mismatches = 634 Gaps = 190 Conservative Substitutions = 0 280 290 300 310 320 330 340 AATTAGACCTGTAGTCCCGACTCTCGCAGGCTGTGTCAACAGCAGGGGTCAGTCTGGTGGAGTCTGT CTCAGAGCGGGTCCGCGCATGA X 10 20 350 360 370 380 390 400 410 CTGGC--GGGGTCTGGCTG--GGGTCTGCCTAGATAGTCTCTGATTTGCTGGGCT--CCATCTGCTG--AAT GTCCCATACGAGACTTGGTGGAGGAGCTTCTGG--AGCAGCTGGAGGCCCGGAGCTACCGGAGGGGCAAT 30 40 50 60 70 80 90				

420 430 440 450 460 470 480
---AGCTGGAGTGTGACATCAATTTGGGGAGCATGTGCTGTCAATCTGGAGGATCATGCTCTGACTAGCTGGGG
100 110 120 130 140 150
CCTCGCTCGTGA-GTTCAAGGACATTAAGGCCCGCTCAAGT-GGCCCTGG-----AAGTCTCAAGGCTGTGTGT
490 500 510 520 530 540
TCCA-TCCATCAGACACATCAAGATGAGATCTGGGC---GGCTGGCT---GGGATGT--CAGGATTCCCATGTG
160 170 180 190 200 210 220 230
TCCCAATTAAGCCGCGAGTCGGCTTGGGAACAGCAACAAAGACCGCTACAAAGATGTGTGTAGCATATCATG-A
240 250 260 270 280 290 300
CAGCAAGAGTATCATCTTTTC--CTGTCTCCAAGAGGAGGACATGGAGATTACAT--CAATGCCAACT--TCATC
310 320 330 340 350
CGG--GGCATAGAT--GGAAAGCCAGGCCCTACATTGGCAGCGCAAGGACCCCTGCCCTCA-CA----CACTGTGT
360 370 380 390 400 410 420
TATATCTGTCTAGTAATCTGACTAACTTAAGG-----GTGCCACCTGGCTGTCCAGCTAGCAGAGATGCTGTGT
430 440 450 460 470 480 490
GACATCTGTGGCCCTGGTTTG-GGAGTTTGGGTCAAGGTAAATCCT-GATGSCCTGTCAAGAGACA-GAATAAT
500 510 520 530 540 550 560
GGCGGAGGAAGTGTGAACGCTACTTGGGCCCGGAGGAGGACCTCTAAAGCTG--GGCCTTTCGCATCA
570 580 590 600 610 620
CTCTGGCTGGGGATCCTTTGTCCACCTGTGGAC-TGTATCTC-TGAAGCTTTTGGCTGTCTGCCAGGTGG
630 640 650 660 670 680 690
CCCTGTGACAAAGAGACACACATGAA--TGCAGACATCACTCTCAGGACCTTCCAGGTTACATTTCCA-GRAGG
700 710 720 730 740 750 760
TGTGCCATCTGTCTGTCCCTATCTGGAGGTCTTTTATCGSTCTG-CCTGGGGTTAAATCTGCTTGGCTGT
770 780 790 800 810 820
AATTCC-GCT-CTGTGACCC-AGCTACA-GTATATGTCTTGGCCAGACACACGGGT--TC-CC--AGCAGT
830 840 850 860 870 880 890
GGTATATGCTTGGCTTAAATCTGCTCGCTGGATGGCGAAGTGTCCATCCACTTCTCTG--TGTCGTCCCTCGAGTG
900 910 920 930 940 950 960
TCTGATCACATTTCTCA-CCATG-GTGGAGAGGCCCGCTGCTCCAGGGGTGGACCTGGACCTGCACCTCTGTGT
970 980 990 1000 1010 1020 1030
TCCAAGGTGAG-GGAGCT-CAGGTGCTGAGGATGGCATGTGCTGTG-TGGTAGGAGA-GCAGT--CTCCGG
1040 1050 1060 1070 1080 1090
TCCACTGCAGTGTGCTCGGACGACACAGG-TGTCCT-GTGGCTGTGTACTATGTGAGGCACTGTGCTGTCT
1100 1110 1120 1130 1140 1150 1160
GAAGCTCATCTGTCCCTTCTCCCTTCTCA-TAACT--CA-GGGGTATGCCCTTCCCTGTGATCTTTCTTTCTTGA
1170 1180 1190 1200 1210 1220 1230
GACCCAGA-CAATCCC-TCCCAACTTCAGTCTCTTCCAAGTGTGCTCGGATGCGGGAACAGC-GGCCTGG
1240 1250 1260 1270 1280 1290
AGC-----AGTGCAGACAGAGG-CAGTACAGGTTCTCT--GTACCACACAGTGGCTC-AGC-TATTTCTCCC
840 850 860 870 880 890
1240 1250 1260 1270 1280 1290

5. US-08-877-150B-36 (1-458)
u49853p_2

Initial Score	-	11	Optimized Score	-	159	Significance	-	-0.41
Residue Identity	-	16%	Matches	-	62	Mismatches	-	250
Gaps	-	42	Conservative Substitutions	-	16			

60 KNRYKDVLPDQTVILSLQEGHSDYINGNFIKVGDSLAIATQGFLPHTLLDFWRLWEFGVIVLMA
70 80 90 100 110 120 130
SERVAAXVAIPTXWGAS--WSS
x x

140 150 160 170 180 190 200
CREJENKRRERYAQEEPLQ-TGLFCITLKEKWLNEDIMLRTLVTFQKESRSVYQLOYNSWDRGVP
||| ||| ||| ||| ||| |||
WRPCTTGRCSSVLSATLRPAQWPGSLKVCVPLKPAVGLGTRTRTATKTMXXHMRQESSFFPCSKRRDMEIT

SSPDHLMWVEARRRGGSPPLCV-HCSACCGTGVLCTVDYVQRLLTQMTDPDFSLFVVYLMKRRQRP
 210 220 230 240 250 260 270
 | | | | | | |
 SMTSSGAXMEARPTURRKDP--CLTICHTWTSAGWFGSLGSRSXSNWPVKRQKMDGGSGVINATGPGCSR-SLNRLL
 100 110 120 130 140 150 160

AAVQTEEOEFLYHTVAONECSTLQNASP-HYONIKENCAPLYDDALFRT--PQALLATP----RPPGGVLK
:
GLSNASQPQR---QHMQ---TSLGSPRLSHRRNSALCTS-YSICPGQTGTGPVALITSPWWRPAAS-K

340
 350
 360
 370
 380
 390
 400
 410
 SISVPGSGHAMADTYAEQKKGAGAGNGSTGTCTGTCARSAAEPLYSKVTPRAQRCAHAEDARGTLPG
 : : : : : : : : : : : : : : :
 GLDLDFSVSTAVLA--ADBQ-----VSCALLTMXGCCXPROSLPTVSYSKWRCNSGLQCQRQSSTGS
 230 240 250 260 270 280

	420	430	440	450	X
RVPA--	--DQSPAGS	GAVDVAGS	GAQTGG	LGFLN	RTGRPKGPRDP--P-----AEWTRV
		:	:		:
CITQWL	SYSPALSR	TPATTR	SRRTALQ--	SARKPS	SGPGQPCLPHPHGQEGFSGASRCLRP
10	300	310	320	330	340 350 360

LWCRSVALBRAQGRGRPPARTPRTARWLHVPSPDRW
370 380 390

6. US-08-877-150B-36.(1-458)
1149853p_5

```
Initial Score      = 9  Optimized Score = 114  Significance = -0.42
Residue Identity  = 13% Matches      = 33  Mismatches   = 188
Gaps              = 13  Conservative Substitutions = 17
```

170 180 190 200 210 220 230
KWLNEDIMLRTLAVTQKESRSYQIQYMSWFDGRVPSSPDHMLANWEARRLOGSGPEPLCVHCSAGCRT
|||
FFFFFCPEKLLFXIRPTSPTH
X X 10 20

240
GVLC TVDYVQLLLTOMIPDPSLEDFVLKMRKQPAAVQTEEQRYFYHTVQAFMFCSTQNASPHYQNTKE
250
260
270
280
290
300
TXHTAQOYGGSRSPA--ITENQAITGTALVTPVSTLLEDPGALWAFPCASXSLALQSIVLHLLLRHQARK
30
40
50
60
70
80
90

NCAPLYDDALFLRTPQALLAIIPRPFGVLRISVSPSP - GHAMADTYAEQBKRGAGAGSGTGTCTG-T-G
:
SFGQELGAVPLSPVRP-PCVPDPVAGHVEPPCCSKGCCFWNAPAALPALCPKPRIASAPQRKCOPWGGSGAQA

100	110	120	130	140	150	160
380	390	400	410	420	430	440
ARSAEAPLSKVTPRAQPCG-AHAEDARTGLGRRVPADQSPAG--SGAYEDVAGGA-QTGGLGFNLRIGRPR	:	:	:	:	:	:
170	180	190	200	210	220	230
PRCSXPELLVAGMHWAGLRXSGRRPLPCR--LERSSPXSGSGGWCPCGCGRIAPELGCTGTCTAPLSALLQ	:	:	:	:	:	:
450	X					
KGRDPDPAEWTRV	:	:	:	:	:	:
240	X	250	260	270	280	290
AAVSASGEPGLGRDQXWEGLSGSAATASHSQORTGHLFVRSQHCSGHRGVQVQALGSGGPPPPW	:	:	:	:	:	: